IBM Data Science Capstone Project report Week 4

Problem Description

When looking to buy a house, location of the house and the cost are two of the major factors to be considered. Location plays an important role as one finds a house with the location having the maximum amenities along with a suitable price value.

In this project I would classify various localities of the city of New Delhi, India based on the amenities offered and the average cost of the houses.

I will use K-means clustering algorithm on the dataset to cluster localities based on the amenities (restaurants, malls, hospitals, etc) within 500 m of circular radius.

Description of Data

The Data will be collected from various webpages using BeautifulSoup and FourSquare api will be used to explore the localities.